European Review of Economic History, **19**, 335–358
doi:10.1093/ereh/hev014 Advance Access publication July 16, 2015

Market potential and regional economic growth in Spain (1860–1930)

JULIO MARTINEZ-GALARRAGA*,
DANIEL A. TIRADO* AND
RAFAEL GONZÁLEZ-VAL**
**Departament D'Anàlisi Econòmica, Facultat D'Economia, Universitat de València, Spain; julio.martinez-galarraga@uv.es; Daniel.Tirado@uv.es*
***Departamento de Análisis Económico, Facultad de Economía y Empresa, Universidad de Zaragoza and Institut D'Economia de Barcelona (IEB), Spain; rafaelg@unizar.es*

In this study, parametric and nonparametric techniques are employed to analyze the effect of changes in regional market potential on the growth of Spanish regions during the period 1860–1930. The study of Spain sheds light on whether the construction of new transport infrastructures coupled with changes in trade policy ultimately shaped regional growth trajectories. The study draws upon new evidence of per capita gross domestic product for Spanish provinces and combines this evidence with Harris' market-potential function to measure regional market potential. Results show that market potential had a positive influence on regional economic growth, particularly between 1900 and 1930.

1. Introduction

Regional income inequality is a recurring feature of apparently well-integrated economies such as that of the European Union. As pointed out by Puga (2002), in 1995 nearly a quarter of European citizens lived in regions with a per capita gross domestic product (GDP) that was below 75 percent of the European average. Arguably, full integration across Europe is still a work in progress, and these differences between European countries will tend to diminish in the long term. Yet, the magnitude and persistence of regional imbalances are still of great concern for national economies that are often considered the embodiment of long-lasting political and economic integration. The study of the trends and determinants of regional income inequality during past processes of national market integration could be of great use. It could help identify the main forces at work, and may shed light on the evolution of regional inequality in modern instances of economic and political integration such as the European Union.

From a theoretical point of view, international and regional economics have explained income disparities in terms of differences between regions' endowments of natural resources, factors of production, infrastructure, and technology (Barro and Sala-i-Martin 1991). Simply removing obstacles that hinder the flow of goods and/or factors would cause the convergence of factor returns and living standards. Yet, as stated in the new economic geography (NEG) literature, forces such as agglomeration economies, which are overlooked in conventional analysis, can affect regional disparities—even without large differences in the underlying characteristics of the regions—and prevent convergence.

Developments in endogenous growth theory and NEG have provided an economic foundation to explain the relation between market size and regional growth. These two bodies of theory posit that market size fosters growth because greater market size allows for exploitation of increasing returns in knowledge creation from investment in R&D, in human capital training, and in activities such as manufacturing or services (Baldwin and Martin 2004). In such circumstances, market integration favors regional specialization and leads to larger markets in activities characterized by the presence of increasing returns, thereby driving regional growth.

From an empirical point of view, economic history studies have shown that economic growth following the integration of several regions can initially lead to an increase in regional per capita income disparities. Williamson (1965) provided evidence for this assertion by examining the long-term evolution of regional inequality in the USA. He first posited that regional inequality within national economies would follow an inverted *U*-curve throughout the process of economic growth. This would entail growing inequality during the nineteenth century and convergence from then on. He concluded that structural change and industrial specialization were responsible for the increase in inequality during the initial stages of economic growth.

Kim (1998) studied in detail the determinants of long-term industrial specialization across the United States, and concluded that relative factor endowments and scale economies acted as the main determinants. In a similar vein, Combes *et al.* (2011) studied the long-term evolution of economic disparities between French départements, concluding that the concentration of the spatial distribution of manufacturing and services traced an inverted *U*-curve starting in the mid-nineteenth century. Interestingly, in line with the arguments proposed in the NEG literature, they found that the existence of agglomeration economies would be a relevant factor for understanding regional income evolution in France between 1860 and 1930.

Turning to Spain, Rosés *et al.* (2010) showed that a long phase of growing regional disparities accompanied the early stages of the Spanish economy's integration. In these early years (1860–1930), the emergence of large differences in production structures across regions favored the upswing in regional economic inequality. In particular, these authors showed that, between 1860 and 1930, regions specialized in industrial production enjoyed higher levels of per capita income[1]. Martinez-Galarraga (2012) showed that industrial specialization of a small number of Spain's regions was determined, among other factors, by regional differences in market potential.

In summary, from the theoretical and empirical literature reviewed in the previous paragraphs, a hypothesis emerges to explain the rise in regional economic inequality during the initial phases of economic development. In a context of economic integration and reductions in transport costs, differences in regional market size would have favored the industrial specialization of a small group of regions. This specialization would have helped these regions to benefit from increasing returns in manufacturing, and would have let them achieve higher rates of economic growth, thereby creating greater economic inequality.

The aim of the current study is to test this hypothesis empirically by analyzing the proximate causes of regional income growth within an empirical framework that captures relevant economic factors. Starting with an NEG model, Ottaviano and Pinelli (2006) derived an empirical strategy based on the estimation of growth regressions. Along with the variables typically found in the traditional growth literature, market potential was also included as a key variable to explain disparities in regional economic growth rates.[2] In the present study, we employ

[1] These authors explored the presence of agglomeration economies using an enlarged version of the Theil Index. In their research, the "between-inequality" component was associated with differences in regions' productive structures, while the "within-inequality" component would potentially capture productivity differentials that would emerge in agglomeration economies.

[2] They analyzed Finnish regions for the period 1977–2002, considering the impact on regional economies of the external shock of the neighboring Soviet Union's collapse.

this methodology to study the determinants of regional inequality in Spain during the period 1860–1930. We hereby present this example as an illustrative case study.

Analysis of Spanish market integration and economic growth offers a three-fold contribution to the body of existing research. The first contribution of this research is that it entails cross-regional analysis, which yields two basic advantages. One of these advantages relates to the analysis of growth trajectories of territories that share a common institutional framework. In contrast to cross-country studies, this commonality reduces the varying impact on growth resulting from institutional differences (Redding 2010). Furthermore, as claimed by Head and Mayer (2006), regional analysis offers scholars a better grasp of the effects of agglomeration economies because interregional transport costs tend to be lower than international costs.

The second contribution of our research is its time span, namely the period from 1860 to 1930. These years span a period of Spanish history characterized by an acceleration in the integration of regional markets owing to enormous advances in national transport networks. Numerous studies have shown that within the increasingly integrated economy of this period market access acted as a determinant of industry location (Martinez-Galarraga 2012), intensity and direction of migration flows (Pons *et al.* 2007), and regional wage levels in the manufacturing sector (Tirado *et al.* 2013). Hence, the period seems to be an especially suitable context in which to analyze the impact of market potential on regional economic disparities, which are measured at the aggregate level using recent provincial GDP estimates by Rosés *et al.* (2010).

Because of new transport networks and changes in Spanish trade policy, the integration of markets asymmetrically altered regional market potential across Spain. The asymmetric nature of changes in regional market size derives from two main sources. First, investment in infrastructure, despite favoring overall reduction in transport costs, had an asymmetric effect on regional economic centrality. In particular, investment in infrastructure led to growth in Madrid's economic centrality, which acted as a major rail transport hub for Spain (Tirado *et al.* 2013). Second, at the end of the nineteenth century, Spain and its trade partners adopted protectionist trade policies that reduced the geographical advantages enjoyed by coastal regions.

This asymmetrical impact endows the Spanish case with a particularly appealing feature. As Redding (2010) pointed out, the analysis of the effects of market access on the remuneration of factors faces an important empirical hurdle: it is difficult to disentangle the effects of market access from other determinants of comparative economic development such as locational fundamentals. This means that the relationships emerging from a significant number of empirical analyses suffer from the problem of causal indeterminacy. Economists have advocated the analysis of this type of relation in the context of exogenous changes in the relative market size of regions or territories. This would then involve analyzing the impact of these changes on factor returns to validate the causal nature of the relationships. Davis and Weinstein (2002), Wolf (2007), and Redding and Sturm (2008) provide examples of such an approach. The study of Spain during this period therefore offers a case study that allows us to analyze whether a new transport infrastructure, as well as changes in trade policy that affected Spanish regions' relative market potential, ultimately shaped regional growth trajectories.

The third contribution of the study resides in the construction of the measure of market potential. In the absence of data on bilateral trade flows between territories (Redding and Venables 2004), NEG scholars have commonly used Harris' market-potential function to measure regional accessibility. This measure is based on the assumption that regional income levels and the geodesic distances between regions are representative of regional market access. Especially, when applied to history, however, the consideration of bilateral transport costs instead of geodesic distances in the calculation of the domestic market potential yields some clear advantages. For instance, it allows us to include different transport modes like railways or coastal shipping, analyze by mode of transport the exact routes used in the transportation of commodities, and consider the freight rates applied by companies. All of these additional factors can be studied, considering that their evolution over time may vary for a number of reasons. Numerous factors may affect transport and trade costs. These factors include the country's terrain and geography, the emergence of new transport technologies, the level of investment in transport infrastructures, the design of the network (often politically decided), and even trade policy. Because Spain's progress in transport infrastructure and changes in Spain's trade policies asymmetrically altered the economic distance between regions, it seems advisable to use a measure of the regional market potential that accounts for these changes. We therefore employ Harris' market-potential function to measure regional market potential, which calculates the economic distance between territories according to changes in transport networks, variations in transport costs, and the tariff policy in force at the time.

The remainder of the article is structured as follows. First, we offer a brief summary of the historical process of market integration and economic growth of the Spanish economy. In doing so, we present new evidence on long-term regional inequality patterns based on recent estimations of per capita GDP for NUTS 3 Spanish regions (provinces) between 1860 and 1930 (Rosés *et al.* 2010). Section 3 explains the proposed measure of regional market potential. In Section 4, we present the results from a nonparametric analysis of the relation between market potential and regional economic growth. Section 5 presents results from the empirical (parametric) analysis, and discusses the main findings. Finally, Section 6 sets forth the conclusions of our study. Appendix A describes sources of data employed in the analysis.

2. Market integration, industrial location, and regional inequality in Spain (1860–1930)

From a historical point of view, major advances occurred during the integration of national markets and industrialization in the nineteenth century. The reduction in trade costs within countries was linked to the elimination of institutional obstacles that hindered the free movement of goods and factors between regions. Furthermore, the fall in transport costs during the Industrial Revolution also led to a drop in trade costs. In the case of Spain, the integration of the domestic market ran in parallel with an increase in the spatial concentration of manufacturing and with a rise in regional income inequality. In this section, we review the main evidence regarding these issues.

Economic integration of Spain's regional economies took place during the second half of the nineteenth century. Before then, during the *Antiguo Regimen* (Ancient Regime), the Spanish market was fragmented, and consisted of a collection of largely unconnected local and regional markets. Historians have stressed two key reasons for this. First, the persistence of institutional obstacles hindered interregional trade. These included the use of different currencies, weights, and measures across regions, and the existence of internal tariffs. Second, deficiencies in Spain's transport infrastructure had stymied development in Spain, a country that has had to confront serious geographical obstacles throughout its history. Yet, the second half of the nineteenth century was witness to a progressive integration of the domestic market, thanks to the institutional reforms undertaken by a sequence of liberal governments and progress in the transport system.

First, nineteenth-century political reforms strengthened property rights and reduced transaction costs that interfered in economic relations and impeded the free movement of goods within Spain. Importantly, reforms eliminated the main restrictions on trade (including tariffs and domestic customs), suppressed guilds and the *Mesta* (a medieval association of cattle farmers), disentailed land (*desamortización*), abolished entailed states (*mayorazgos*), and unified the system of weights and measures that had hitherto varied from region to region. In addition, improvements in the transport sector proved to be a determining factor for the integration of the Spanish market. The introduction of the railway and advances in other modes of transport (road and coastal shipping) triggered a fall in transport costs.

The outcome was the gradual integration of the national market for goods for the main traded products. This integration was characterized by convergence in regional prices. Various studies have demonstrated the gradual convergence of regional grain prices from the beginning of the eighteenth century until its culmination in the second half of the nineteenth century (Peña and Sánchez-Albornoz 1983). In addition, the integration of capital and labor markets led to massive advances. In the case of capital markets, the main events that affected the monetary and banking systems favored a reduction in interest rate differentials across regions (Castañeda and Tafunell 1993). Lastly, Spain's labor market integration has also been extensively analyzed. Rosés and Sánchez-Alonso (2004) showed that PPP-adjusted rural and urban wages converged across different locations prior to World War I, despite low rates of internal migration.

From 1869 onwards, this context of internal market integration was accompanied by a progressive economic openness towards neighboring countries. The reduction in tariff protection levels was most notable at the end of the 1880s, when Spain signed several trade treaties with its main trading partners. Nevertheless, the last decade of the century witnessed an important change in terms of the Spanish economy's integration with external markets. In 1883, the gold convertibility of the peseta was abandoned, thus debilitating Spain's position in the international capital markets. Furthermore, from 1892 onwards, the return to protectionism seriously threatened external integration. Thus, an inverted *U*-curve plots the evolution of Spanish international trade during this period. The rate of openness of the Spanish economy increased during the second half of the nineteenth century. This trend, however, began to reverse in the 1890s (see figure 1).

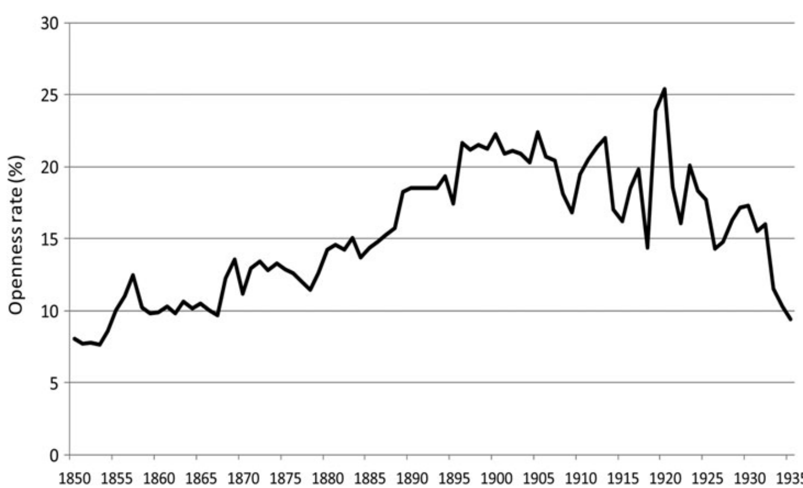

Figure 1. *Openness rates (percent). Spain, 1850–1935.*
Source: *1850–1913: data for exports and imports from Prados de la Escosura (2010); 1914–1935, Tena (2005); GDP figures from Prados de la Escosura (2003).*

Advances in the integration of Spanish national markets for goods and factors drove an intense process of regional specialization. From the middle of the nineteenth century to the outbreak of the Civil War (1936–1939), industrial production gradually agglomerated in a small number of provinces, a development well documented by historiographers (Nadal 1987). A recent estimation of regional gross value added (GVA) in Spanish industry (Tirado and Martinez-Galarraga 2008) allows to show how geographical concentration of manufacturing activities evolved over the period studied (figure 2). The general pattern is one of an increase in the concentration of industry across Spain's provinces up to the 1930s. This is a similar dynamic to that of other countries like the United States (Kim 1995) and France (Combes *et al.* 2011). These nations also experienced an upsurge of agglomeration of industrial production during the early stages of national market integration and industrialization.

Economic historians have investigated the roots and causes of this notable increase in the spatial concentration of manufacturing before the Spanish Civil War. They have done so by studying the role played by the two major explanatory theories from the literature: traditional trade theory (comparative advantage in a Heckscher–Ohlin setting) and NEG. Rosés (2003) found evidence that the *home market effect* was driving early Catalan industrialization (around the 1860s). Tirado *et al.* (2002), in line with Kim (1995), identified scale economies and market size as the determinants of Spain's industrial geography in the mid-nineteenth century. By the end of the century, the explanatory power of these NEG effects had increased in parallel with advances in the economic integration process. Recently, Martinez-Galarraga (2012), adopting the approach developed by Midelfart-Knarvik *et al.* (2002), confirmed and extended the previous findings by Tirado *et al.* (2002). As the domestic market became integrated and industrialization progressed in Spain during the second half of the nineteenth century, NEG forces grew to be the main determinant of Spain's industrial landscape. In particular, although comparative advantage factors were a feature, the scale effects suggested by Krugman (1991) played a decisive role: industries with increasing returns tended to concentrate in provinces with better access to demand up to the 1930s.

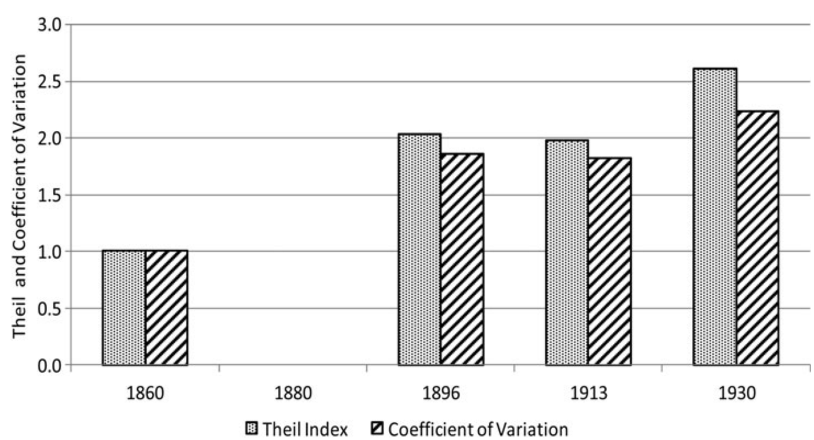

Figure 2. *Inequality in the distribution of industrial GVA in Spain at province level, 1860–1930 (1860 = 1).*
Source: *Tirado and Martinez-Galarraga (2008).*

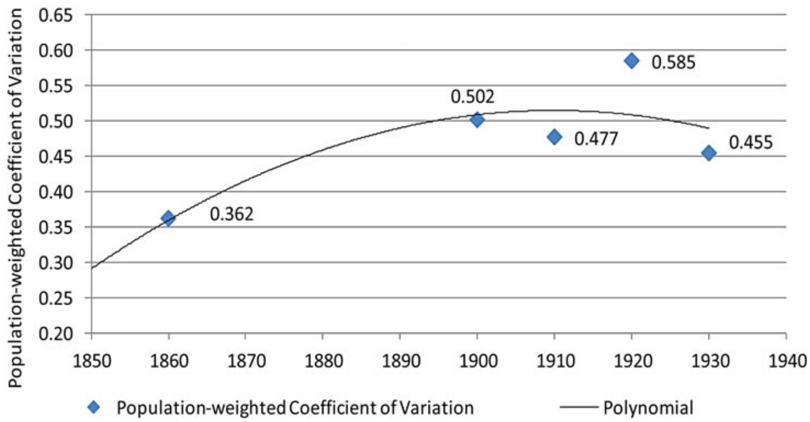

Figure 3. *Regional per capita GDP inequality, Spanish NUTS 3 (1860–1930).*
Source: *Rosés et al. (2010)*.

In the context of the Spanish economy, researchers have also tested the existence of higher wages in regions that have greater market-potential resulting from the agglomeration of manufacturers in core regions (backward linkages), and the attraction of these wages for generating migratory flows of worke*rs* (forward linkages). First, following Hanson's (1998, 2005) influential research based on the Krugman wage equation, Tirado *et al.* (2013) examined the existence of a spatial structure in industrial nominal wages in 1920s Spain. The results verify that wages were higher in regions with greater market potential, and the authors prove the existence of a wage gradient centered in Barcelona, the main industrial center in interwar Spain. This work also confirmed that domestic market potential became more relevant as the Spanish economy and main European markets increased protectionism during the 1920s, while the wage gradient centered on Barcelona declined.[3] Second, following Crozet (2004), Pons *et al.* (2007) established a direct relationship between migration decisions and the market potential of host regions during the 1920s, thus verifying the presence of forward linkages in internal migrations between Spain's provinces in the interwar years.

After describing the evidence gathered at the manufacturing sector level, the next step is to focus on regional inequality in terms of per capita GDP (Rosés *et al.* 2010). Taking the population-weighted coefficient of variation as a measure of inequality, figure 3 plots the long-term evolution of regional income per capita disparities at the province level (NUTS 3). In Spain, the second half of the nineteenth century witnessed a remarkable increase in regional income inequality. Then, in the early decades of the twentieth century, this process came to a halt. Accordingly, figure 3 shows a tendency towards the stabilization of income per capita inequality.

What are the determinants of this evolution? The evidence presented so far shows that NEG forces were driving the agglomeration process observed in the manufacturing sector in Spain from the mid-nineteenth century to the outbreak of the Spanish Civil War. But was the impact of geography limited to the manufacturing sector? Did NEG-type mechanisms have an effect on income per capita? In other words, did market potential and its evolution have an

[3] This exercise thus contributed to the existing theoretical and empirical NEG debate about the effects of international integration on the internal geography of countries (Krugman and Livas 1996; Crozet and Koenig 2004).

impact on regional inequality during the early stages of economic growth in Spain? In order to answer these questions, we must construct a sound measure of regional accessibility. The next section is thus devoted to building a market-potential indicator.

3. Measuring regional market potential in Spain, 1860–1930

In NEG multi-regional models, the capacity of different locations to attract firms and workers varies according to their position relative to one another. Although in NEG studies Harris' market-potential function has tended to rely on geodesic distances between locations, there are fundamental reasons to consider bilateral transport costs, especially in historical studies.[4] During the period analyzed in this article, remarkable changes occurred in transport technologies. These changes include the expansion of railways and steam navigation. In addition, as a peninsula, Spain's geography made it possible to transport commodities between provinces by both land and sea (coastal shipping). Traditionally, inland transport had transported goods by road, and had been expensive because of the country's mountainous topography and the poor state of roads. Furthermore, the absence of navigable rivers deprived Spain of an alternative, cheaper form of transport. The construction of the railway network triggered a reduction in transport costs, but its expansion was gradual and therefore some regions were able to benefit from railway transportation earlier than others. Furthermore, the policymakers responsible for designing the railway network decided to create a radial network with its hub in Madrid, the geographical center of the country. At the same time, coastal shipping underwent some key advances, such as the transition in navigation from sail to steam and improvements in port facilities. Overall, these changes often follow a regionally asymmetric pattern, thereby unevenly affecting regional transport costs and accessibility.

In order to analyze the potential relationship between market potential and regional income growth, we gathered two different types of empirical evidence. First, we used the new estimations of regional GDP constructed in Rosés *et al.* (2010). Second, we adopted a new estimation of the market potential of Spanish regions for the five benchmarks considered. The proposed regional market-potential measure came from the so-called nominal market potential or Harris' market-potential function, defined as:[5]

$$\mathrm{MP}_r = \sum \frac{M_s}{d_{rs}},$$

where M_s is a measure of the economic size of province[6] s (i.e., GDP) and d_{rs} is the distance. (In this case, d is equal to the bilateral transport costs between provinces r and s.)

[4] The measurement of transport costs has been and remains the subject of much debate. The geodesic straight-line distance, the real distance as a function of the available infrastructure, the distance measured in time, and the transport costs that include the distances and the freight rates, are the alternatives used in empirical studies. A review of the literature from an NEG perspective can be found in Lafourcade and Thisse (2008).

[5] This measure of market access could be considered an ad hoc indicator because it has been neither built upon a solid theoretical foundation nor derived from a structural estimate. Advances made by NEG models, however, have helped overcome the lack of theoretical foundation for this empirical measure of market potential. Particularly, Combes *et al.* (2008) derived an expression for the real market potential (RMP) that establishes a relationship with Harris' (1954) equation. Nonetheless, from an empirical perspective, when compared with structural estimates of the market potential, Head and Mayer (2004) expressed a preference for the performance of Harris' equation.

[6] Provinces are Spanish NUTS 3 regions. The insular territories (Balearic and Canary Islands) have not been included, giving a total of 47 provinces.

Drawing on this expression, Martinez-Galarraga (2014) computed a measure of Spanish NUTS 3 market potential for the years 1867, 1900, 1910, 1920, and 1930, based on the study by Crafts (2005).[7] Market-potential figures were obtained as follows. First, the market potential of a Spanish province r was disaggregated into two components: the domestic market potential (DMP_r), to which each province's self-potential (SP_r) was incorporated, and the foreign market potential (FMP_r) between the provincial and the international nodes. Hence, the market potential of a province r (MP_r) was calculated as the sum of the domestic and foreign market potential:

$$\mathrm{MP}_r = \mathrm{DMP}_r + \mathrm{FMP}_r.$$

According to this expression, the domestic market potential for each of the 47 provinces r can be calculated as follows

$$\mathrm{DMP}_r = \sum_{1}^{s=46} \frac{M_s}{d_{rs}} + \mathrm{SP}_r,$$

where

$$\mathrm{SP}_r = \frac{M_r}{d_{rr}}$$

is the measure of the self-potential of each province r, and d_{rr} is calculated by taking a distance θ_{rr} equivalent to a third of the radius of a circle with an area equal to that of the province (Crafts 2005):

$$\theta_{rr} = \frac{1}{3}\sqrt{\frac{(\text{Area of the province}_r)}{\pi}}$$

The next expression yields the foreign market potential of province r (FMP_r):

$$\mathrm{FMP}_r = \sum_{1}^{f=4} \frac{M_f}{d_{rp}} \cdot \text{distance}_{pf}^{\delta} \cdot \text{tariff}_f^{\gamma},$$

where $d_{rp} = 1$ if r is a coastal province, and $d_{rp} = d_{rs}$ if r is an inland province. In this case, M_f is the economic size of the foreign market; d_{rp} captures the transport costs from the inland provincial node to the nearest Spanish sea port p; distance_{pf} is the distance between the Spanish sea port and the international node f; tariff f_f are the mean tariffs applied in the foreign country f; and δ and γ are the elasticities estimated by international trade gravity equations associated to the coefficients for distance and tariffs, respectively.

Hence, the total market potential of province r (MP_r) is the sum of the following terms, the first two corresponding to the domestic market potential (including the self-potential of province r) and the third corresponding to the foreign market potential:

$$\mathrm{MP}_r = \sum_{1}^{s=46} \frac{M_s}{d_{rs}} + \mathrm{SP}_r + \left[\sum_{1}^{f=4} \frac{M_f}{d_{rp}} \cdot \text{distance}_{pf}^{\delta} \cdot \text{tariff}_f^{\gamma}\right], \tag{1}$$

where d_{rp} captures whether province r is coastal or inland.

[7] See Martinez-Galarraga (2014) for a detailed description.

The economic size of the provincial markets was measured by the aggregate income. Data on nominal GDP at a NUTS 3 level were obtained from Rosés *et al.* (2010). To measure d_{rs}, we considered transport costs, so data on distances and average transport rates for commodities are needed. Internal transport was assumed to be by railway and coastal shipping. For railway distances, the sources were *Ministerio de Obras Públicas* (1902), and Wais (1987). For distances between ports, electronic atlases provided information on the length of sea journeys.[8] With regard to transport costs, data on railway rates came from Herranz (2005). Coastal shipping rates were obtained from Nadal (1975). This information refers to the transport of coal from Asturias (northern Spain) to eleven peninsular ports. Finally, because coastal provinces could use both rail and coastal shipping to transport goods, volumes of commodities transported by each of these modes between the 1860s and the 1930s were relevant. The relative weight of each transport mode in the coastal provinces came from Frax (1981).

The construction of the measure for foreign market potential was based on the gravity equation for international trade estimated by Estevadeordal *et al.* (2003). The elasticities obtained for distance and tariffs (−0.8 and −1.0, respectively) were used to reduce the size of foreign markets. Nominal GDP of the main trading partners for Spain (France, United Kingdom, Germany and United States) came from Crafts (2005), based on the estimates of Prados de la Escosura (2000). Prevailing exchange rates were applied to convert GDP figures from pounds to pesetas. Maritime distances were again obtained from an electronic atlas, and, finally, tariffs came from O'Rourke (2000) and Mitchell (1998a, 1998b).

Figure 4 allows us to examine the geographical pattern of regional accessibility and its evolution from 1860 to 1930. Throughout the period of study, Barcelona emerged as the province with the greatest market potential. Maps are therefore expressed relative to this province. The evidence shows that the most significant changes in the relative accessibility of the Spanish provinces occurred in the second half of the nineteenth century in parallel with the integration of the domestic market. There was a centrifugal tendency, and the geographical structure evolved towards a clear division between inland and coastal provinces with the latter showing a higher market potential than their inland counterparts, the sole exception being Madrid. Arguably, the expansion of the railway network—all province capitals were connected to the railway network by 1901—could account for a large share of the changes described in the pattern of market potential.[9] Once this dual structure was established at the end of the nineteenth century, the division between inland and coastal provinces persisted throughout the first few decades of the twentieth century.

4. Market potential and economic growth in Spain, 1860–1930: nonparametric evidence

From the nonparametric evidence, we seek to analyze whether the changes in relative market potential of the Spanish regions acted as an explanatory element of regional economic inequality. Ioannides and Overman (2004) have highlighted the advantages of the nonparametric approach over the standard parametric one (e.g., a correlation index). Principally, nonparametric tools do not impose any structure on underlying relationships, which may be nonlinear

[8] www.dataloy.com and www.distances.com.

[9] Access to foreign markets was important to allow coastal provinces to better their position, although self-potential in industrial provinces grew significantly over time. For a further description of the main patterns in provinces' market potential and its components, both domestic and foreign, see Martinez-Galarraga (2014).

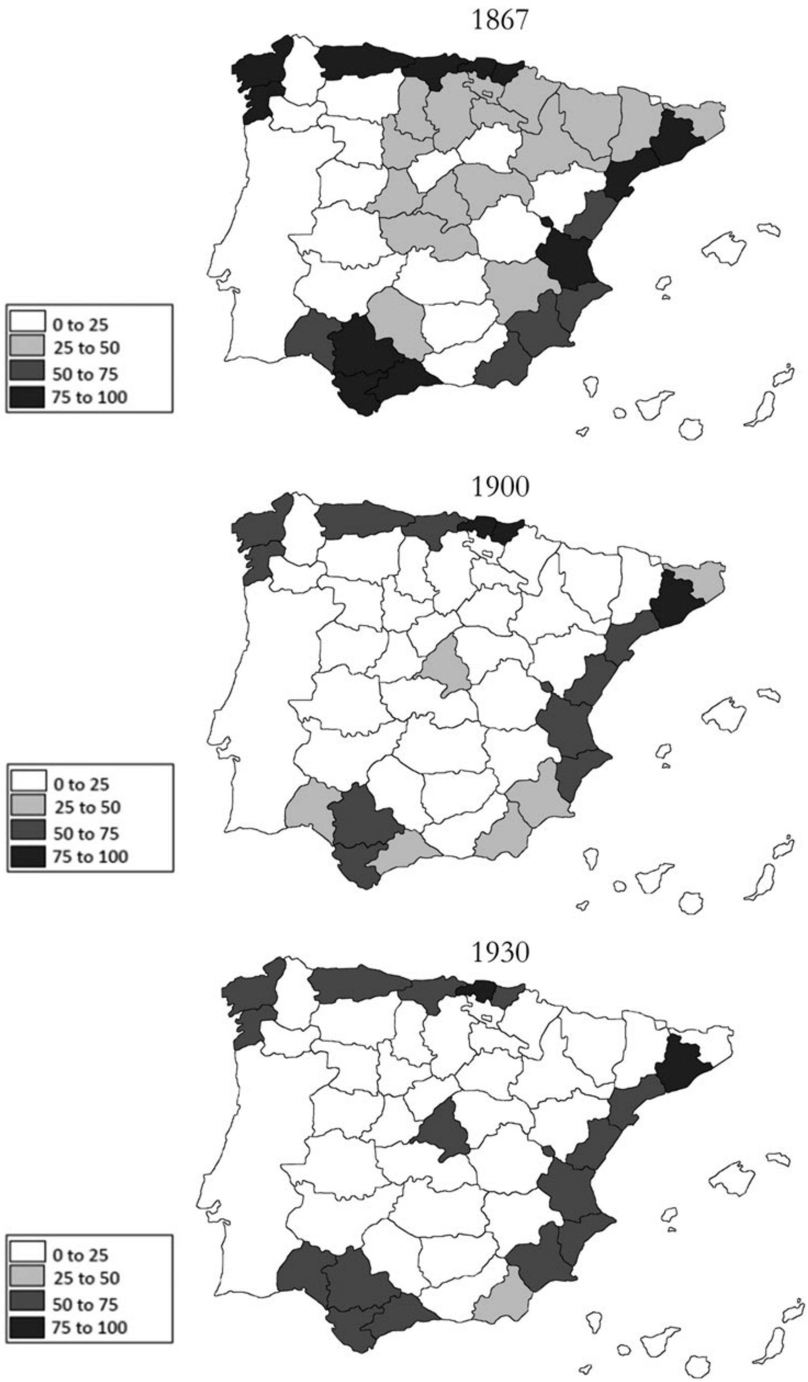

Figure 4. *Market potential in Spanish provinces, 1867–1930 (Barcelona = 100).*
Source: *Martinez-Galarraga (2014).*

(thereby allowing for the nonlinear effects predicted by some theoretical models) and may change over time (no need to restrict the relationship to be stationary, see Ioannides and Overman 2004). For example, the standard correlation index and parametric regressions give us only an aggregate average relationship between income growth and market potential, with the restriction that it must hold over the entire support of the distribution of market potentials. In contrast, the nonparametric estimate lets income growth vary with market potential over the entire distribution. As a first step in the analysis, we present the nonparametric evidence on the relationship between regional per capita GDPs and market potential. To begin with, figure 5 shows the evidence regarding the geographical distribution of regional inequality at three focal points: 1860, 1900, and 1930.

The evidence shown in figure 5 apparently illustrates the presence of a relationship between the relative market access of regions and the corresponding regional per capita GDP levels. The centrifugal pattern observed is similar to the evolution of regional market potential (figure 4), although in the case of per capita GDP the division between inland and coastal provinces is not so marked. To enhance the analysis of these hypotheses, we test the relationship between regional market potential and per capita GDP. To do so, we examine the distribution of regional market potential and the distribution of per capita GDP at the same date. We then study how they are related. Figure 6 shows the stochastic kernel estimations of the distribution of regional market potential, conditional on the distribution of per capita GDP at the beginning and at the end of our period of study. In order to make the interpretation easier, the contour plots are also shown. In 1860, both distributions were clearly independent, and regions with similar levels of per capita GDP had very different values of market potential. By 1930, however, this relationship had changed, becoming positive: regions with high per capita GDP also had high market potential. This result illustrates the emergence of a significant positive relationship between market potential and regional per capita GDP at the end of the period analyzed (i.e., 1930).

Given this change in the relationship between market potential and per capita GDP, we would expect to find a similar relationship between market potential and per capita GDP growth rates. Figure 7 shows the stochastic kernel estimation of the distribution of regional market potential conditional on the subsequent per capita GDP growth rates. Indeed, results point in the same direction: initial market potential and per capita GDP growth between 1860 and 1900 were independent for most of the distribution, while a clear positive relationship emerged for the period 1900–1930. Thus, a higher market potential implied a higher GDP growth rate in 1930, but not in 1860. Overall, these figures indicate a sharp change in the relationship between market potential and per capita GDP over time, from independence to a positive influence of market potential on GDP, especially, in data corresponding to the period 1900–1930, and particularly in the years 1920–1930.[10]

Next, we conduct a nonparametric estimation of the effects of market potential on regional per capita GDP growth. To do this, we estimate the nonlinear relationship between initial market potential and growth using a local polynomial smoothing for the two main subperiods in our sample (1860–1900 and 1900–1930).[11] Figure 8 shows the results, including the 95 percent confidence intervals. These graphs complement figure 7. In the 1860–1900 period,

[10] Figures for all the intermediate periods between 1860 and 1930, omitted from the article due to restrictions on length, are available from the authors upon request.

[11] The local polynomial provides a smoother fits for the growth rate $g_{it} = (\ln pc\text{GDP}_{it} - \ln pc\text{GDP}_{it-1})$ to a polynomial form of $(\ln \text{MP}_{it-1})$ via locally weighted least squares. We used the *lpolyci* command in STATA with the following options: local mean smoothing, a Gaussian kernel function to calculate the locally weighted polynomial regression, and a bandwidth determined using Silverman's rule-of-thumb.

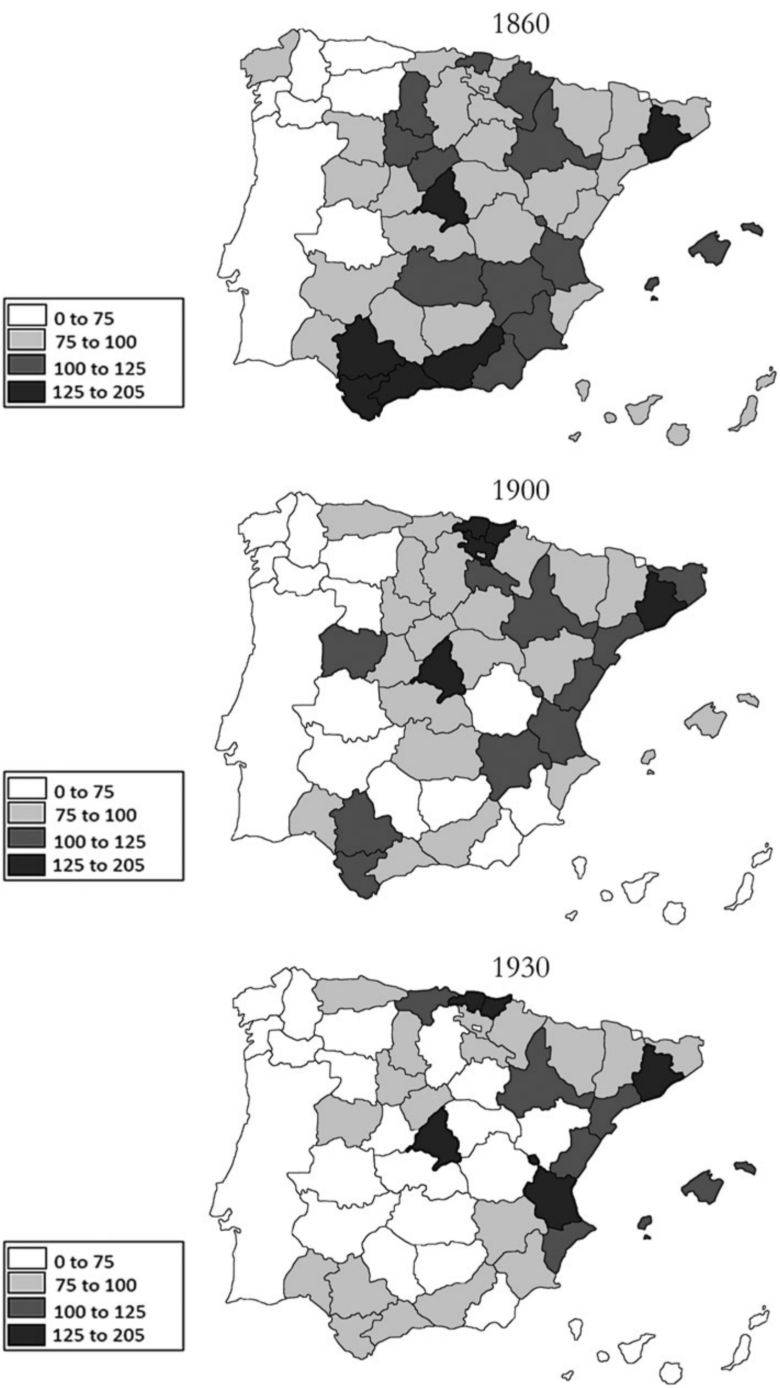

Figure 5. *Per capita GDP in Spanish provinces, 1860–1930 (Spain = 100)*.
Source: *Rosés et al. (2010)*.

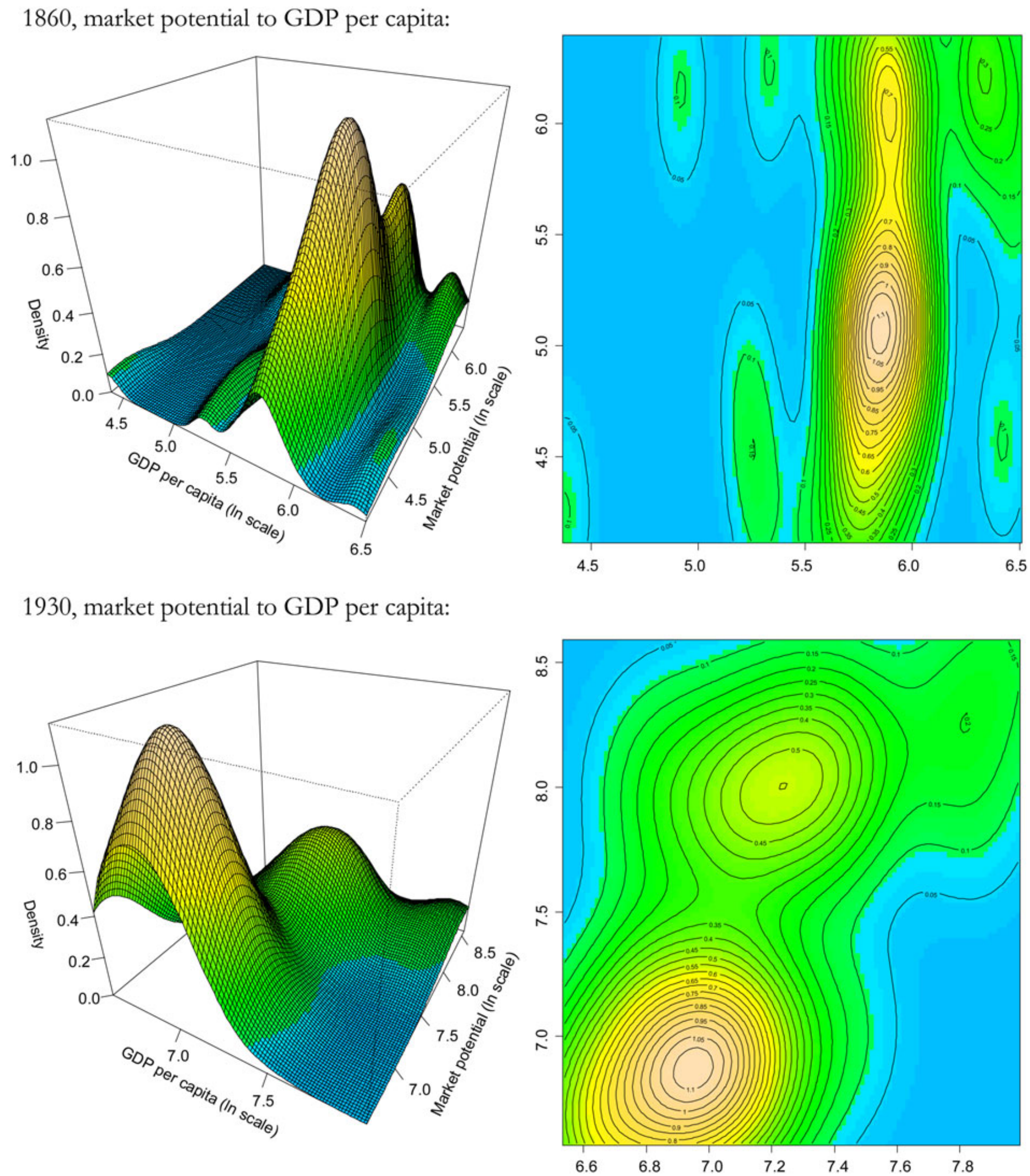

Figure 6. *Stochastic kernel estimates of the relationship between regional market potential and per capita GDP.*

growth can be approximated as a flat line around the value 0.3 for most of the initial market potentials. The relationship is positive only for the highest market potentials, but figure 7 indicates that the density (i.e., number) of regions with the highest market potential was low. Therefore, although a positive relationship between market potential and regional per capita income growth emerged in the two periods under study, there was a temporal evolution implying the increasing influence of market potential over time. Particularly, when focusing on

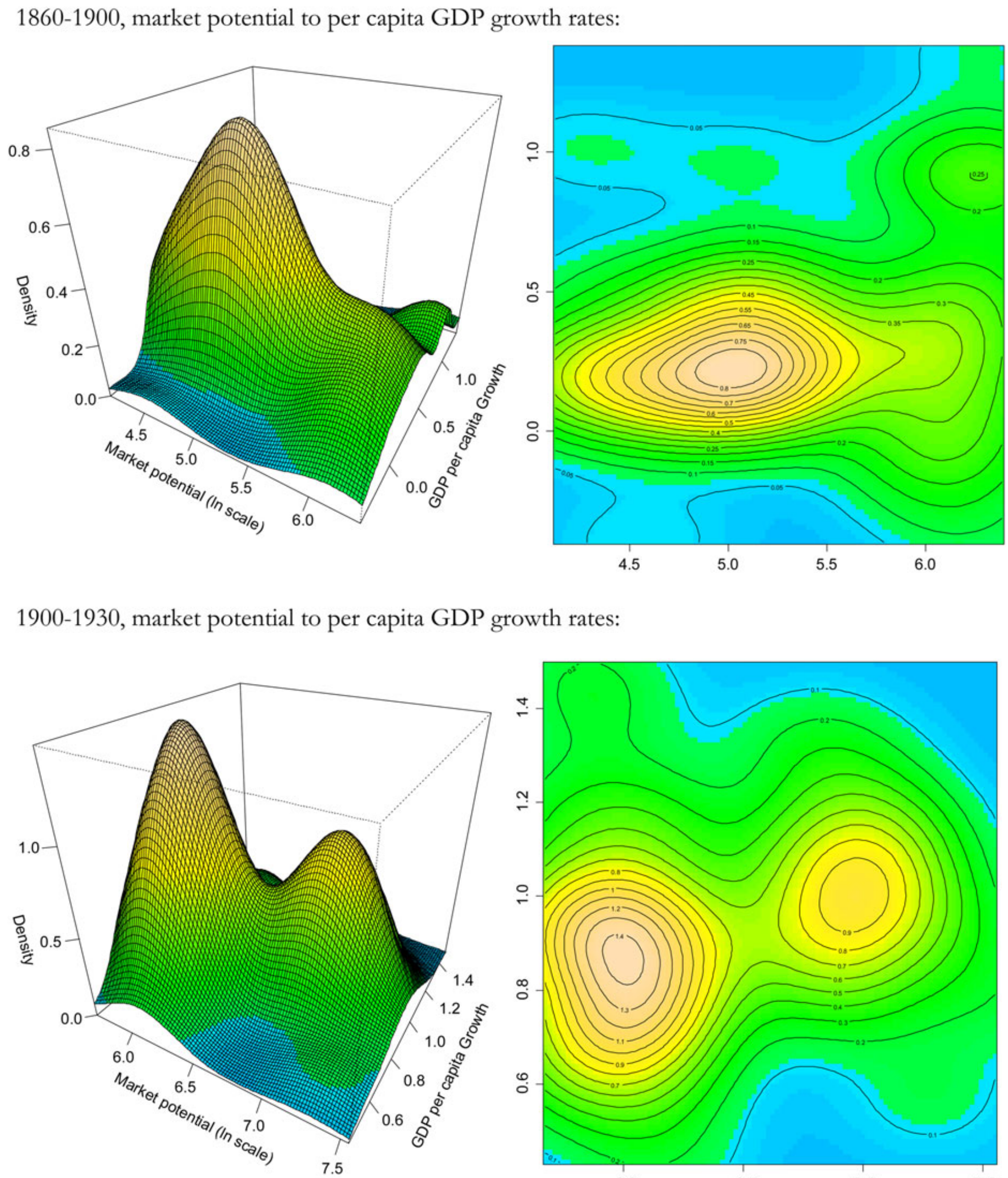

Figure 7. *Stochastic kernel estimates of the relationship between regional market potential and per capita GDP growth.*

regions with a low initial market potential, the effect on mean per capita GDP growth ranged from 0.2 to 0.4 between 1860 and 1900, and from 0.8 to 0.9 between 1900 and 1930. A similar pattern arose for regions with a high initial market potential, although the effect on regional economic growth tended to be higher in these high market potential regions. It ranged from 0.4 to 0.6 in 1860–1900 and from 0.8 to 1 in the period 1900–1930. Moreover, figure 7

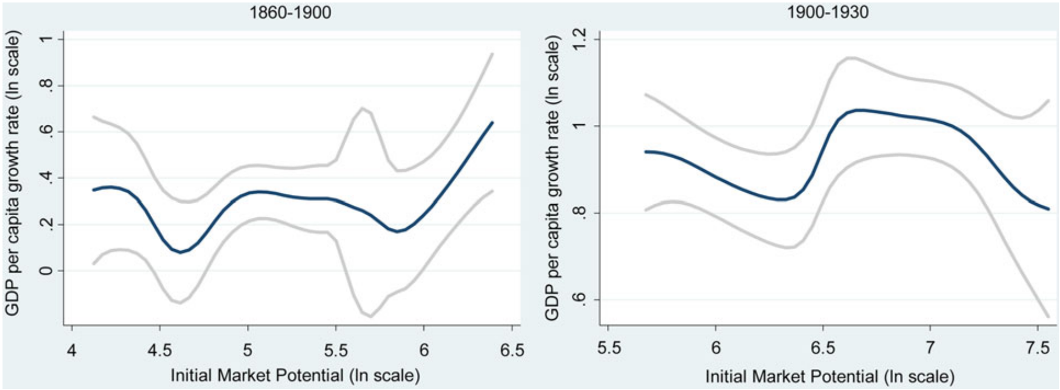

Figure 8. *Nonparametric estimation of the relationship between regional market potential and per capita GDP growth.*
Note: *Nonparametric estimates (local polynomial smoothing) and the 95 percent confidence intervals.*

shows that the density (i.e., number) of regions with the highest market potential also increased substantially over time.

Based on this nonparametric evidence, it is possible to identify the existence of a relationship between regional market potential and regional inequalities. This relationship was more important between 1900 and 1930, once relative regional market potentials had shifted because of the construction of the railway network, the integration of Spain's domestic market, and changes in external tariff policies at the end of the nineteenth century. Hence, because the early decades of the twentieth century constitute the key period in the relationship, the rest of the article presents the results of analysis that aim to probe into this relationship by making use of the theoretical and empirical method proposed by Ottaviano and Pinelli (2006).

5. Empirical analysis

In the parametric analysis, we exploit the panel structure of our data for the years where there was a strong relationship between regional market potential and per capita GDP growth rates. We therefore study the period 1900–1930 using panel data for the subperiods 1900–1910, 1910–1920, and 1920–1930. We begin by estimating standard growth regressions for a set of explanatory variables including a measure of market access. The baseline equation, which resembles that proposed by Ottaviano and Pinelli (2006), takes the following form:

$$\ln(w_t) - \ln(w_{t-1}) = \alpha + \beta \ln(w_{t-1}) + \gamma \ln(\text{access}_{t-1}) + \delta \ln(\text{controls}_{t-1}) + \varepsilon_t, \qquad (2)$$

where the independent variable (i.e., the measure of regional economic performance equal to the logarithmic growth rate of per capita GDP at the province level) is regressed on a set of explanatory variables consistently employed in the growth literature. Notably, and in contrast to cross-country studies, because regions in the same country tend to share the same institutional framework, this exercise does not include a set of institutional variables. Among the explanatory variables, three sets of variables traditionally considered in the growth literature are included. The first set of variables refers to the proximate sources of growth (the stocks of physical capital, human capital, knowledge capital, and infrastructures). The inclusion of these variables (in stocks) controls for the initial endowment of production factors that relate to capital in the production function and thus positively affect regional growth trajectories. Second, the model includes structural change variables (GVA in mining and the regional share of manufacturing in total employment). The pace of structural change can affect regional per capita GDP levels throughout the process of economic integration and industrialization. An alternative convergence mechanism is structural change. The reallocation of productive factors from agriculture to manufacturing tends to increase average productivity given that output per worker is typically higher in manufacturing. In poorer regions, and especially during the early stages of development, agricultural sectors are relatively large and thus the flow of resources towards manufacturing activities may contribute to convergence. We control for this effect in the empirical analysis by including variables that capture regions' productive structure. Finally, we incorporate second-nature geography or NEG variables (market access). We also include regional fixed effects to control for other regional characteristics not accounted for in the specification (e.g., first-nature causes and geography) and region-specific time trends to capture the particular behavior over time of regions in our panel. The dataset and sources appear in Appendix A.

The main explanatory variable in our analysis is regional market potential. In this case, the cross-sectional measure of market potential is normalized by the contemporaneous average market potential to prevent effects from later periods overpowering earlier ones on account of absolute growth in market potential (Black and Henderson 2003). Regional relative market potential (mp_{it}) can therefore be defined as

$$\mathrm{mp}_{it} = \frac{\mathrm{MP}_{it}}{\frac{1}{n_t}\sum_{1}^{n_t}\mathrm{MP}_{jt}}.$$

We use alternative measures of market potential, corresponding to the different components of the market potential (see equation 1): total, domestic, and foreign market potential. We also use a measure that excludes each province's self-potential to reduce some endogeneity concerns (more on this below).

First, we estimate equation 2 by OLS, while correcting for heteroskedasticity using White's method. Nevertheless, an important component of Harris' market-potential function is the contribution of its own GDP to the potential of region i, also known as self-potential. By construction, the explanatory variable (i.e., market potential) and the dependent variable (i.e., per capita GDP growth) therefore influence each other and could be simultaneously determined. Furthermore, because we consider the infrastructures a key element to explaining the changes in the market potential for regions, our main concern relates to the role of these infrastructures. Policymakers tend to improve infrastructures in the most developed regions, but these infrastructures (roads, railways, etc.) undoubtedly also increase the market access of these locations (Holl 2012), generating endogeneity and problems with our specification.

To deal with these two issues, we proceed as follows. First, in some estimations we use a measure of market potential that excludes each region's self-potential. By doing so, changes in regional infrastructures may have affected per capita GDP growth in region i, but we exclude the possible effect of infrastructures on the market potential of region i. In addition, purging self-potential avoids possible simultaneity problems. Second, to tackle the potential endogeneity problem, we re-estimate equation 2 using instrumental variables (IV). We thus need to instrument the market potential variable in the first-stage regressions of the IV estimation.

We use two instruments: a measure of distances across regions and the lagged regional population density. As in Head and Mayer (2006), we use region i centrality, measured as $\sum_{j \neq i}^{n-1} d_{ij}^{-1}$. Distances are the same as in the calculation of market potential. An alternative strategy would be to consider the distance to the nearest main industrial center (Madrid, Barcelona, or Bilbao)[12], but our choice of centrality index is a more flexible measure because it does not explicitly impose a center. Population can serve as a good measure of market potential, and some studies have employed it instead of GDP (Black and Henderson 2003; Ioannides and Overman 2004). To be cautious, we use the lagged value of the regional population density. Values from 1860 are hence used to estimate market potential in 1900, and so on. Finally, region-specific time trends, which we also include in our specification, capture variation over time in the instrumented market-potential measure in the first-stage regression.

Table 1 shows the results of the OLS estimation of equation 2. The first column corresponds to an unconditional convergence regression. The estimated coefficient is clearly significant and negative, indicating convergence across Spanish regions. In the rest of the columns, convergence is stronger when all controls are added. In addition, only two of the four measures of relative market potential are significant: total market potential and domestic market potential. Estimated coefficients are positive in all cases. As explained previously, however, these OLS estimations are not robust, so we instrument the market-potential variables using the lagged regional population density and the centrality measure, and we estimate the second-stage regressions by 2SLS. We use the 2SLS estimator instead of the GMM because the GMM estimator might have poor small sample properties (Baum *et al.* 2003) and thus might not perform well on a data like the current one (i.e., with 141 observations). Table 2 reports these results. Table 2 also shows some statistics from the first-stage regressions. Our instruments seem to perform well: R^2 in the first-stage regressions exceeds 0.85 in all specifications, the weak instruments hypothesis is always rejected using the Stock–Yogo test, and almost all models pass the overidentification test (Hansen $\mathcal{J}$ statistic) for any significance level. The foreign market-potential model is the only exception, with the null hypothesis rejected at 5 percent but not at 1 percent.[13]

With regard to the control variables, we first focus on the proximate sources of growth. All coefficients associated with stock of knowledge capital—proxied by number of patents per capita—and provincial stock of infrastructures are significant and have the expected signs. That is, they confirm the presence of a positive relationship between the relative stocks of these cumulative factors and regional growth. However, human capital stock—proxied by provincial literacy rates—is not significant in any model. The structural change variable, GVA in mining, relates negatively to regional economic growth in most of the regressions. The share of manufacturing in total employment (another structural change variable) also shows, as expected, a negative sign. This result would be capturing the effect of reallocation of resources from agricultural to manufacturing activities, as previously explained. Yet, the coefficient is only significant in the specification that includes the domestic market potential.

The IV results confirm the positive effect of initial market potential on per capita GDP growth. The estimated coefficients of the four measures of market potential are significant and positive,

[12] We also estimate the equation using the distance to the nearest main industrial center instead of the centrality measure. Results were qualitatively similar.

[13] The complete results of the reduced regressions, first-stage regressions, and all the tests, excluded from the article due to restrictions on length, are available from the authors on request.

Table 1. *Regional growth regressions (OLS)*

Variables	(1)	(2)	(3)	(4)	(5)
Initial GDP per capita	−0.28***	−1.37***	−1.36***	−1.34***	−1.37***
Literacy rate		0.08	0.13	−0.20	0.14
Number of patents per capita		0.42***	0.44***	0.57***	0.43***
GVA in mining		−0.02***	−0.02***	−0.01	−0.02***
Share of manufacturing in total employment		−0.86	−0.39	−3.06***	−0.40
Total stock of infrastructures		0.32***	0.39***	0.19**	0.38***
Relative market potential		0.15**			
Relative market potential without self-potential			0.03		
Relative domestic market potential				0.45***	
Relative foreign market potential					0.03
Regional fixed-effects	No	Yes	Yes	Yes	Yes
Region time	No	Yes	Yes	Yes	Yes
R^2	0.14	0.79	0.77	0.82	0.77
Observations	141	141	141	141	141

Note: Dependent variable: per capita GDP growth rate (ln scale). All variables in logarithmic scale, except rates and relative market potentials.
Significant at the *10 percent, **5 percent, and ***1 percent levels. All specifications include a constant.

Table 2. *Regional growth regressions (IV, 2SLS)*

Variables	(1)	(2)	(3)	(4)
Initial GDP per capita	−1.38***	−1.40***	−1.34***	−1.39***
Literacy rate	0.03	0.09	−0.21	0.15
Number of patents per capita	0.40***	0.33***	0.57***	0.36***
GVA in mining	−0.01***	−0.02***	−0.01	−0.02***
Share of manufacturing in total employment	−1.19	−0.15	−3.11**	−0.32
Total stock of infrastructures	0.27***	0.32***	0.18**	0.33***
Relative market potential	0.27***			
Relative market potential without self-potential		0.27**		
Relative domestic market potential			0.46***	
Relative foreign market potential				0.13**
Regional fixed-effects	Yes	Yes	Yes	Yes
Region time	Yes	Yes	Yes	Yes
First-stage, uncentered R^2	0.96	0.94	0.98	0.89
First-stage, F-test (p-value)	38.21 (0.00)	19.97 (0.00)	21.81 (0.00)	28.42 (0.00)
Hansen J statistic, p-value	0.129	0.074	0.175	0.036
Uncentered R^2	0.86	0.83	0.89	0.85
Observations	141	141	141	141

Note: Dependent variable: per capita GDP growth rate (ln scale). Instruments: centrality and lagged regional population density. All variables in logarithmic scale, except rates and relative market potentials.
Significant at the *10 percent, **5 percent, and ***1 percent levels. All specifications include a constant.

and the estimated values are notably different. The estimated value of the coefficient of relative market potential excluding self-potential (0.272) is close to that of relative market potential (0.266). This result is noteworthy. The regression using the market potential that excludes each province's self-potential should be especially robust because, by excluding the self-market of the region, we avoid some potential endogeneity and simultaneity concerns.

The biggest coefficient, however, belongs to the relative domestic market potential (0.458). Interestingly, this coefficient is (more than three times) bigger than that of the foreign market potential, which has the lowest estimated coefficient (0.125). Both are significant. This result must be analyzed in the context of the implementation and reinforcement of a protectionist trade policy by successive Spanish governments from the late nineteenth century until the 1920s, as explained in Section 2. As Spain consolidated its protectionism, domestic market potential became more relevant than foreign markets as a driver of regional per capita growth rates. This result confirms evidence from previous analyses of Spain's industrial sector during the interwar years (Tirado *et al.* 2013).

6. Conclusions

Regional income inequality is prevalent in apparently well-integrated economies such as those of the European Union. In fact, as pointed out in Section 1, although income differences across EU Member States have fallen over the past years, inequalities between regions within each Member State have persisted. So, despite considerable resources having been devoted to reducing this divergence, regional inequality remains a matter of concern for European policymakers. We argue that empirical analysis of the historical determinants of regional income inequality during a period of long-term growth, and internal and external integration of national economies could help to understand differences in economic growth across territories.

From a theoretical point of view, international and regional economics have explained income disparities in terms of differences between regions regarding endowments of natural resources, factors of production, infrastructure, and technology. In this context, removal of obstacles to the flow of goods and/or factors alone would cause convergence of factor returns and living standards. As posited by NEG theory, however, relevant forces, which can affect regional disparities—even without large differences in underlying characteristics—and prevent convergence are overlooked in traditional analysis. NEG theoretical models state that the interaction between transport costs, increasing returns, and market size under a monopolistic competition framework can lead to spatial agglomeration of economic activity and to the upsurge of income differences across regions (Krugman 1991).

To contribute to this empirical debate, we analyzed the determinants of regional inequality in Spain during the period 1860–1930. Spain serves as an illustrative case study for (at least) two reasons. First, in Spain, more than 150 years of economic and political integration have failed to eradicate per capita GDP differences between regions. Second, such long-term analysis examines whether the factors highlighted by NEG models (market access) had a relevant effect on regional economic growth during the early stages of economic growth and integration of the Spanish national market.

To complete these aims, we first used an empirical model that drew upon the work of Ottaviano and Pinelli (2006). Within a single framework, the model analyzed how factors highlighted by Solow-type growth and NEG literature affected long-term regional economic growth in Spain. Second, we used new evidence on regional Spanish per capita income and on market potential for the years 1860–1930. To complete our data set, we also gathered data commonly used in growth regressions to identify the main forces that explain regional growth in Spain between 1860 and 1930.

Overall, the results of the empirical analysis indicate that geography matters when explaining asymmetric regional growth, especially during the period 1900–1930. During the second half of the nineteenth century, agriculture was still the predominant sector in the Spanish economy; industry had begun to take root in only a limited number of regions. Nevertheless, our results show that since the beginning of the twentieth-century NEG forces, through market potential, had a positive influence on provincial growth differentials, even when we controlled for proximate causes of growth. The emergence of agglomeration forces would be the outcome of the interaction between increasing returns to scale and a drop in transport costs. Several major events caused these changes. Completion of the railway network led to a fall in transport costs, propelling the integration of the Spanish economy. Spanish industrialization during the second half of the nineteenth century also provoked major changes. In addition, Spain's protectionist trade policy from the late nineteenth century onwards enhanced the role of the domestic market as an explanatory factor for differences in regional economic growth.

Acknowledgements

This research has received financial support from the projects ECO2012-39169-C03-02, ECO2013-45969-P, and ECO2013-41310-R (Ministerio de Economía y Competitividad), DGA (ADETRE research group), XREPP (Generalitat de Catalunya) and FEDER. Previous versions of this article were presented at the 60th Annual North American Meetings of the Regional Science Association International (Atlanta, 2013), the 54th Congress of the European Regional Science Association (Saint Petersburg, 2014), the XL Reunión de Estudios Regionales (Zaragoza, 2014), and the Economic History Seminars at the Universitat de Barcelona and Universidad de Murcia. We wish to thank the participants, two anonymous referees and the editor for their comments. The usual disclaimer applies.

References

Baldwin, R. and Martin, P. (2004). Agglomeration and regional growth. In J.V. Henderson and J.F. Thisse (eds), *Handbook of Regional and Urban Economics, vol. 4, chapter 60*. North Holland: Elsevier, pp. 2671–711.

Barro, R. and Sala-I-Martin, X. (1991). Convergence across states and regions. *Brookings Papers on Economic Activity* **1**, pp. 107–82.

Baum, C., Schaffer, M. and Stillman, S. (2003). Instrumental variables and GMM: estimation and testing. *Stata Journal* **3**, pp. 1–31.

Black, D. and Henderson, V. (2003). Urban evolution in the USA. *Journal of Economic Geography* **3**, pp. 343–72.

Castañeda, L. and Tafunell, X. (1993). Un nuevo indicador para la historia financiera española: la cotización de las letras de cambio a corto plazo. *Revista de Historia Económica* **11**, pp. 367–83.

Combes, P.P., Mayer, T. and Thisse, J.F. (2008). *Economic Geography. The Integration of Regions and Nations*. Princeton: Princeton University Press.

Combes, P.P., Lafourcade, M., Thisse, J.F. and Toutain, J.C. (2011). The rise and fall of spatial inequalities in France: a long-run perspective. *Explorations in Economic History* **48**, pp. 243–71.

Crafts, N. (2005). Market potential in British regions, 1871-1931. *Regional Studies* **39**, pp. 1159–66.

Crozet, M. (2004). Do migrants follow market potentials? An estimation of a New Economic Geography model. *Journal of Economic Geography* **4**, pp. 439–58.

Crozet, M. and Koenig, P. (2004). Trade liberalization and the internal geography of countries. In T. Mayer and J.L. Mucchielli (eds), *Multinational Firms' Location and the New Economic Geography*. Cheltenham: Edward Elgar, pp. 91–109.

DAVIS, D.R. and WEINSTEIN, D.E. (2002). Bones, bombs, and break points: the geography of economic activity. *American Economic Review* **92**, pp. 1269–89.

ESTEVADEORDAL, A., FRANTZ, B. and TAYLOR, A.M. (2003). The rise and fall of world trade, 1870–1939. *Quarterly Journal of Economics* **118**, pp. 359–407.

FRAX, E. (1981). *Puertos y comercio de cabotaje en España, 1857–1934*. Madrid: Banco de España.

HANSON, G.H. (1998, 2005). Market potential, increasing returns and geographic concentration. *Journal of International Economics* **67**, pp. 1–24.

HARRIS, C.D. (1954). The market as a factor in the localization of industry in the United States. *Annals of the Association of American Geographers* **44**, pp. 315–48.

HEAD, K. and MAYER, T. (2004). The empirics of agglomeration and trade. In J.V. HENDERSON and J.F. THISSE (eds), *Handbook of Regional and Urban Economics*, **vol 4**. North Holland: Elsevier, pp. 2609–69.

HEAD, K. and MAYER, T. (2006). Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics* **36**, pp. 573–94.

HERRANZ, A. (2005). La reducción de los costes de transporte en España (1800–1936). *Cuadernos Económicos del ICE* **70**, pp. 183–203.

HERRANZ, A. (2008). *Infraestructuras y crecimiento económico en España (1850–1935)*. Madrid: Fundación de los Ferrocarriles españoles.

HOLL, A. (2012). Market potential and firm-level productivity in Spain. *Journal of Economic Geography* **12**, pp. 1191–215.

IOANNIDES, Y.M. and OVERMAN, H.G. (2004). Spatial evolution of the US urban system. *Journal of Economic Geography* **4**, pp. 131–56.

KIM, S. (1995). Expansion of markets and the geographic distribution of economic activities: the trends in U.S. regional manufacturing structure, 1860–1987. *Quarterly Journal of Economics* **110**, pp. 881–908.

KIM, S. (1998). Economic integration and convergence: U.S. regions, 1840–1987. *Journal of Economic History* **58**, pp. 659–83.

KRUGMAN, P. (1991). Increasing returns and economic geography. *Journal of Political Economy* **99**, pp. 483–99.

KRUGMAN, P. and LIVAS ELIZONDO, R. (1996). Trade policy and the Third World metropolis. *Journal of Development Economics* **49**, pp. 137–50.

LAFOURCADE, M. and THISSE, J.F. (2008). New economic geography: a guide to transport analysis. *PSE Working Papers*, 2008-02.

MARTINEZ-GALARRAGA, J. (2012). The determinants of industrial location in Spain, 1856–1929. *Explorations in Economic History* **49**, pp. 255–75.

MARTINEZ-GALARRAGA, J. (2014). Market potential estimates in history: a survey of methods and an application to Spain, 1867–1930. *EHES Working* Papers, No. 51.

MIDELFART-KNARVIK, K.H., OVERMAN, H., REDDING, S. and VENABLES, A.J. (2002). The location of European industry. In EUROPEAN COMMISSION. (eds), *European economy. European integration and the functioning of product markets, 2/2002*. Brussels: Directorate-General for economic and financial affairs, pp. 213–69.

Ministerio de Obras Públicas (1902). *Estadísticas de Obras Públicas, años 1899 y 1900*. Madrid: Dirección General de Obras Públicas.

MITCHELL, B.R. (1998a). *International historical statistics. Europe 1750–1993*. London: Macmillan.

MITCHELL, B.R. (1998b). *International historical statistics. The Americas 1750–1993*. London: Macmillan.

NADAL, J. (1975). *El fracaso de la revolución industrial en España, 1814–1913*. Barcelona: Ariel.

NADAL, J. (1987). La industria fabril española en 1900. Una aproximación. In J. NADAL, A. CARRERAS and C. SUDRIÀ (eds), *La economía española en el siglo XX. Una perspectiva histórica*. Barcelona: Ariel, pp. 23–61.

NÚÑEZ, C.E. (1992). *La fuente de la riqueza: educación y desarrollo económico en la España contemporánea*. Madrid: Alianza.

O'ROURKE, K.H. (2000). Tariffs and growth in the late 19th century. *The Economic Journal* **110**, pp. 456–83.

Ottaviano, G.I.P. and Pinelli, D. (2006). Market potential and productivity: evidence from Finnish regions. *Regional Science and Urban Economics* **36**, pp. 636–57.
Peña, D. and Sánchez-Albornoz, N. (1983). *Dependencia dinámica entre precios agrícolas. El trigo en España, 1857–1890. Un estudio empírico*. Madrid: Banco de España.
Pons, J., Paluzie, E., Silvestre, J. and Tirado, D.A. (2007). Testing the New Economic Geography: migrations and industrial agglomerations in Spain. *Journal of Regional Science* **47**, pp. 289–313.
Prados de la Escosura, L. (2000). International comparisons of real product, 1820–1990: an alternative data set. *Explorations in Economic History* **37**, pp. 1–41.
Prados de la Escosura, L. (2003). *El progreso económico de España 1850–2000*. Bilbao: Fundación BBVA.
Prados de la Escosura, L. (2010). Spain's international position, 1850–1913. *Revista de Historia Económica/Journal of Iberian and Latin American Economic History* **28**, pp. 173–215.
Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography* **2**, pp. 373–406.
Redding, S.J. (2010). The empirics of new economic geography. *Journal of Regional Science* **50**, pp. 297–311.
Redding, S.J. and Sturm, D.M. (2008). The costs of remoteness: evidence from German division and reunification. *American Economic Review* **98**, pp. 1766–97.
Redding, S.J. and Venables, A.J. (2004). Economic geography and international inequality. *Journal of International Economics* **62**, pp. 53–82.
Rosés, J.R. (2003). Why isn't the whole of Spain industrialized? New economic geography and early industrialization, 1797–1910. *The Journal of Economic History* **63**, pp. 995–1022.
Rosés, J.R. and Sánchez-Alonso, B. (2004). Regional wage convergence in Spain 1850–1930. *Explorations in Economic History* **41**, pp. 404–25.
Rosés, J.R., Martinez-Galarraga, J. and Tirado, D.A. (2010). The upswing of regional income inequality in Spain (1860–1930). *Explorations in Economic History* **47**, pp. 244–57.
Sáiz, J.P. (2005). Investigación y desarrollo: patentes. In A. Carreras and X. Tafunell (eds), *Estadísticas Históricas de España, siglos XIX–XX*, **vol. II**. Bilbao: Fundación BBVA, pp. 835–72.
Tena, A. (2005). Sector exterior. In A. Carreras and X. Tafunell (eds), *Estadísticas Históricas de España, siglos XIX–XX*, **vol. II**. Bilbao: Fundación BBVA, pp. 573–644.
Tirado, D.A. and Martinez-Galarraga, J. (2008). Una nova estimació retrospectiva del VAB regional industrial. Espanya (1860–1930). *Documents de Treball de la Facultat de Ciències Econòmiques i Empresarials, Col·lecció d'Economia* E08/192.
Tirado, D.A., Paluzie, E. and Pons, J. (2002). Economic integration and industrial location: the case of Spain before World War I. *Journal of Economic Geography* **2**, pp. 343–63.
Tirado, D.A., Pons, J., Paluzie, E. and Martinez-Galarraga, J. (2013). Trade policy and wage gradients: evidence from a protectionist turn. *Cliometrica* **7**, pp. 295–318.
Wais, F. (1987). *Historia de los ferrocarriles españoles*. Madrid: Fundación de los Ferrocarriles Españoles.
Williamson, J.G. (1965). Regional inequality and the process of national development: a description of the patterns. *Economic Development and Cultural Change* **13:4**(Part II), pp. 3–84.
Wolf, N. (2007). Endowments vs. market potential: what explains the relocation of industry after the Polish reunification in 1918? *Explorations in Economic History* **44**, pp. 22–42.

Appendix A: Regression variables: data and sources

A.1 Regional performance measures

Data on Spanish GDP at a NUTS3 level of aggregation between 1860 and 1930 come from Rosés *et al.* (2010). Population by province is obtained from Population Censuses.

A.2 Explanatory variables

A.2.1 Proximate sources of growth

(a) Physical capital. The regressions include the initial level of per capita GDP.
(b) Human capital. The stock of human capital in each province is proxied by data on literacy rates coming from Núñez (1992).
(c) Knowledge capital. The stock of knowledge capital is measured by the number of patents per capita. Unfortunately, only the number of patents registered at a NUTS2 level of aggregation based on the information provided by Sáiz (2005) is available. Therefore, NUTS2 data have been applied to each one of the provinces within each NUTS2 region.
(d) Infrastructures. Information on the total stock of infrastructures is provided by Herranz (2008).

A.2.2 Structural change variables

(e) Gross value added in mining at the province level comes from Rosés *et al.* (2010).
(f) Share of manufacturing in total employment by province comes from Rosés *et al.* (2010).

A.2.3 Market access or second nature geography

(g) Market potential. Alternative measures of regional market potential based on Martinez-Galarraga (2014) are used.